

Ryan Spencer

INF1344

EDA and Visualizations

December 15th, 2022

Looking at the UN data set, I was presented with several questions. What causes these mass migratory patterns? In most cases, they are caused by human intervention, namely wars. But are there any other human interventions that affect migrant stock? Policies or economic strife? As the report will outline, there are a few examples of unique factors that affect migrant stock.

Methods

When looking at the UN Migrant data set, I employed the tenets of *Exploratory Data Analysis* as outlined by John Tukey. Sort, group, box charts, range and median enabled my analysis to find patterns either by instances, greatest or least counts, and within ranges. The importance of just looking at the data first, sorting by largest migrant stock and then grouping by country provided to be a good tactic to understand migratory patterns in the data set

To begin visualizing the UN Migrant dataset, I began by reading *Envisioning Information* by Edward Tufte. Through this process, I'd like to apply his key five principals:

To visualize I applied Edward Tukey's five principal to data ink:

- 1) Maximize the data-ink ratio.
- 2) Erase non-data-ink.
- 3) Erase redundant data-ink.
- 4) Revise and edit.

My takeaway is to create graphs that present a clear and concise answer in the most minimalist way. The use of small multiples and editing for chartjunk are key points that I'd like to cover in this report.

Results

Table 1/DF1 – International Migrant Stock at Midyear 1990-2015

As I will outline individually, each of the tables contain unique observations so a blanket approach to EDA would not suffice. For Table 1, “International Migrant Stock at Midyear 1990-2015” contains several key variables to explore: Year, Sex, Migrant Stock Count, and Country or Major Region. Compared to the other tables, Table 1 offers a cumulative report that spans over several decades. In a sense, a bird’s eye view of overall trends, anomalies, and patterns are goals to explore.

When isolating for year, I came across a huge migrant count for a relatively small country: Russia. Compared to the populations of India and the USA, Russia had an outsized migrant count. Upon researching further and contextualizing to world events. The migration was triggered by the fall of the USSR.

Migrant Stock at

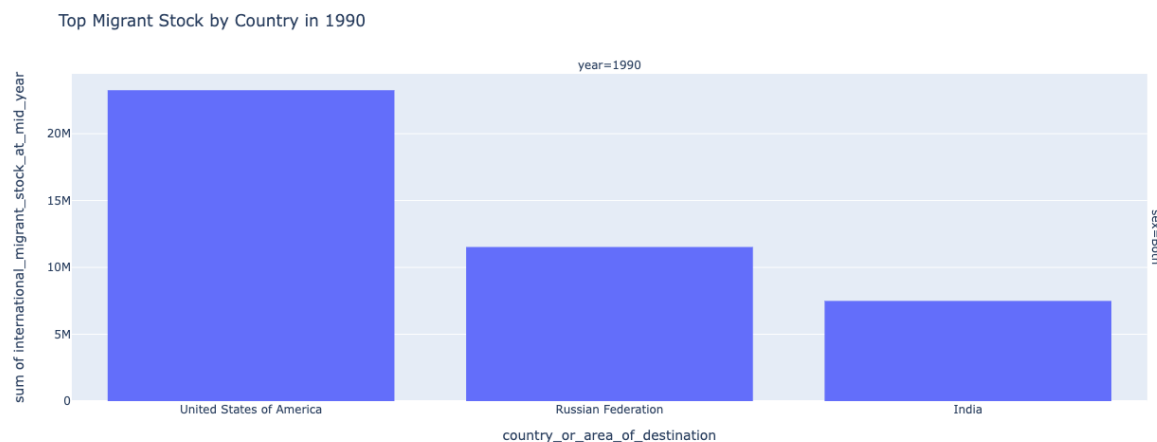


Figure 1. To Migrant Stock by Country in 1990

Simple histogram that conveys the issues at hand, why does Russia have a spike in migrant stock in 1990?

Looking to Tufte’s idea of multiples to compare changes, I created the below histogram to show the rise of migrant stock by Continent and Year (1990-2015). As you can see below in Figure 2,

Europe had a surge of migrant stock that is rapidly growing. Asia, on the other hand, has a greater overall population than Europe so the findings can better be explored in Table 3 when percentage of migrant stock compared to population are introduced.

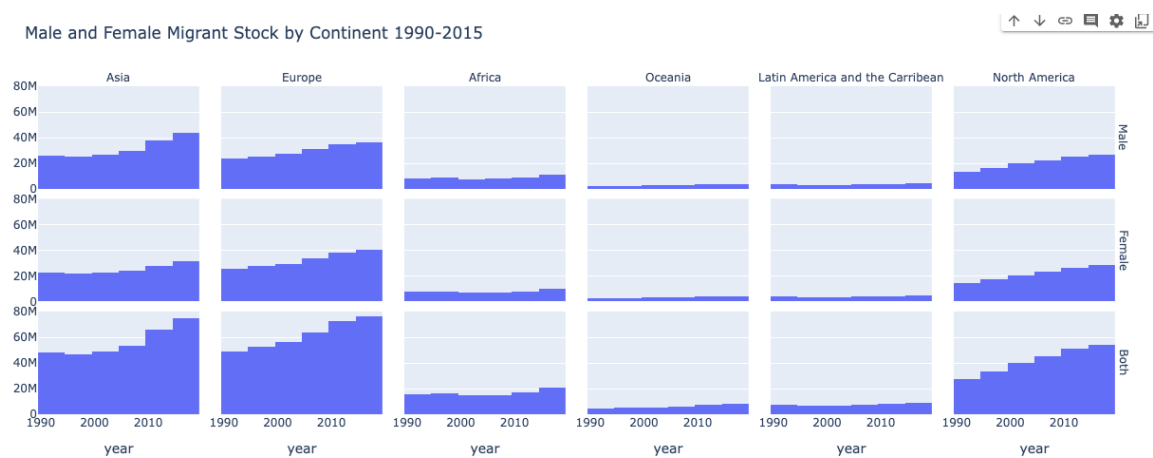


Figure 2. Male and Female Migrant Stock by Continent 1990-2015

UN Migrant Stock

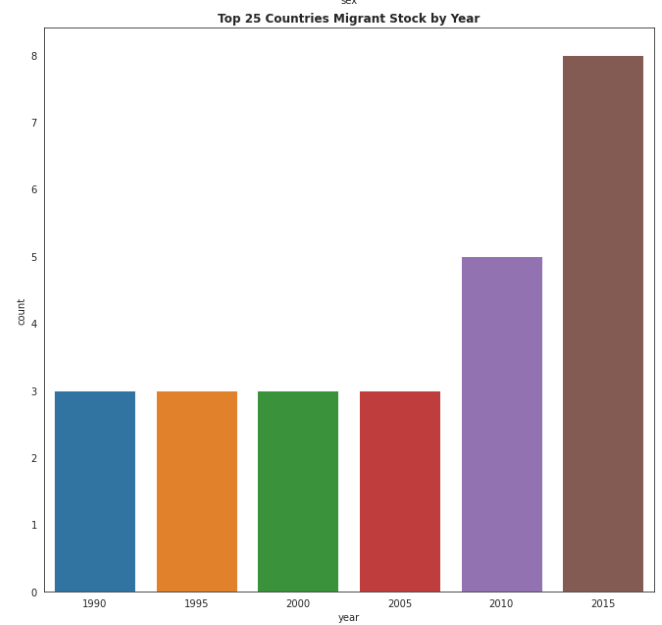
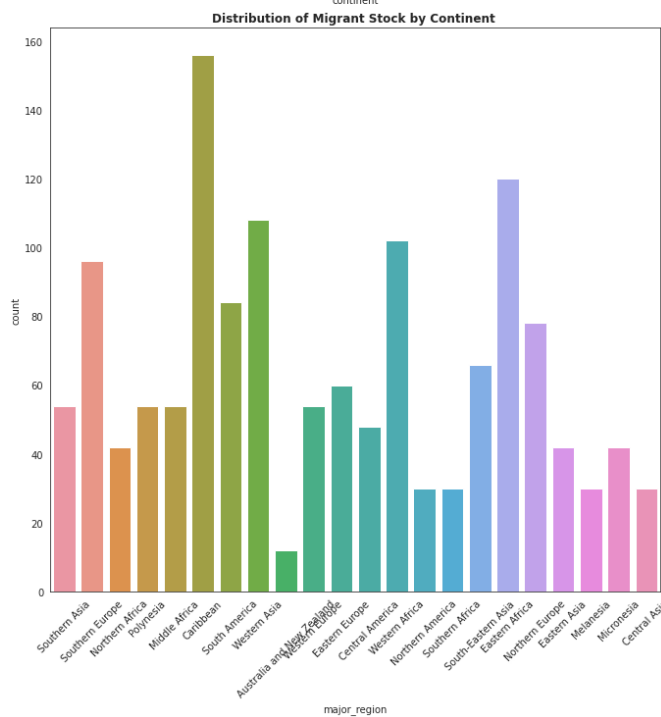
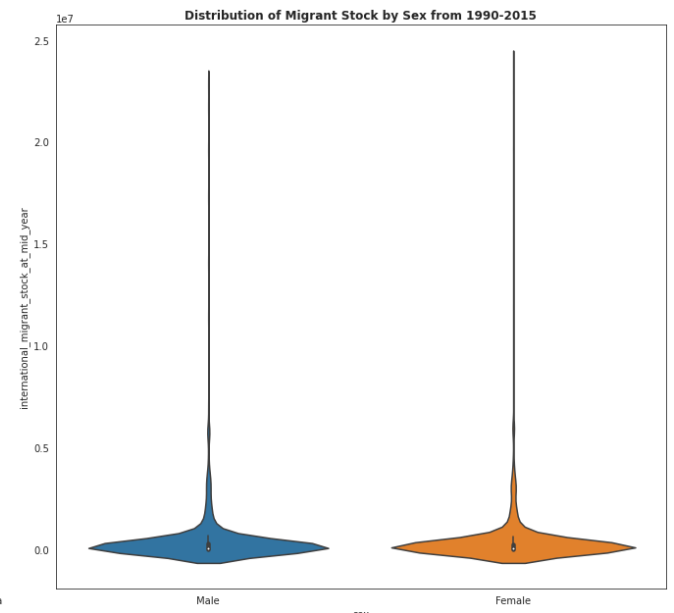
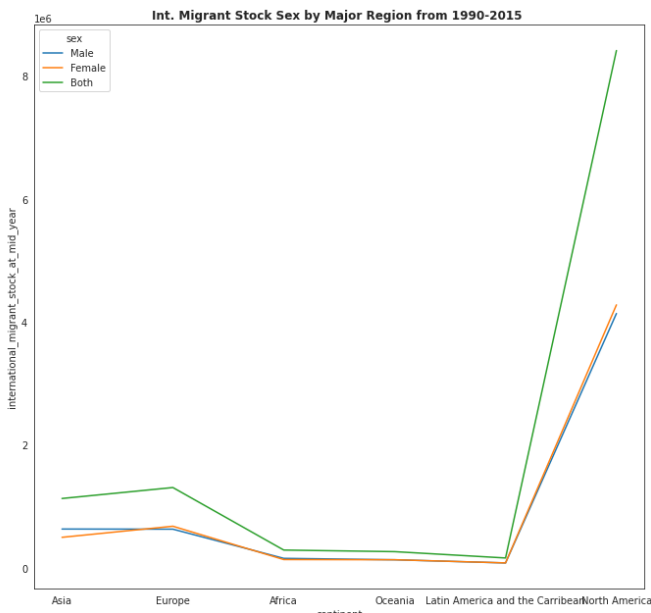


Fig 3a. International Migrant Stock Sex by Major Region from 1990-2015

Fig 3b. Distribution of Migrant Stock by Sex from 1990-2015

Fig 3c. Distribution of Migrant Stock by Continent

Fig 4d. Top 25 Countries Migrant Stock by Year

For Fig 3a, the line graph best displays the findings by giving a cursory overview of the data. Clean lines and colour coded for quick reference, it's an overview to get acquainted with the data. It's a quick skim to see sexes are mostly on parity and to note when they are not. This is a key takeaway when starting to understand migration patterns. The findings show a spike in both sexes for North America and Europe. Female migrants eclipse Male migrants in those continents. I will explore further in Table 4 when analyzing the Percentage of Female Migrant Stock.

Fig 3b. is a violin plot that shows the range and media of sexes in the data. The findings from Fig 3a are confirmed: there are more female migrants as plotted by the whisker. The median of both sexes is similar in the violin plots.

Fig 3c. Distribution of Migrant Stock by Continent 1990-2015

Fig 3c is a histogram displaying Migrant Stock totals by continent. There is a spike in migrant totals in the Caribbean which is curious. I chose histograms as they display count data well and to show better comparisons between continents.

Fig 3d. is a histogram that shows the Migrant Stock total for the Top 25 Countries. The spike in migrant stock for 2015 is a finding. Again, histograms were chosen to show count data and comparison.

Table 2/DF2 – International Migrant Stock in Thousands 1990-2015

This table contains the total populations of countries in a thousand count. This data could prove useful when comparing Migrant Stock from Table 1 by giving a perspective on human migratory patterns. A good reference point as well to know the distribution of populations among countries.

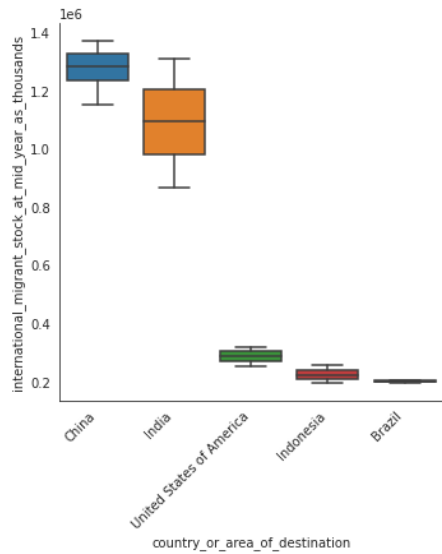
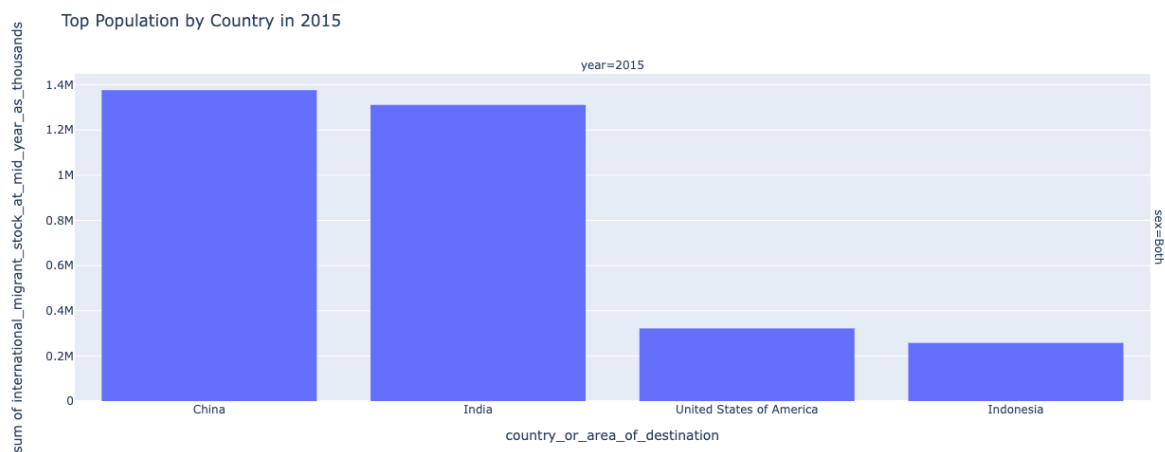


Fig 5 Top 5 Int. Migrant Stock by Country 1990-2015

Fig 5 is a box chart that shows China and India having the largest populations over 1990-2015 period. It's interesting to note that India appears to have a faster growing population as plotted by the top whisker. The boxplot was employed here to show the median as well as the potential population growth of the countries.

Fig 6. Top Migrant Stock by Country in 2015



In Figure 6, we confirm the findings of the box chart and see that India is in fact poised to eclipse China's population. The histogram displays a more nuanced yet expansive finding than the box plot

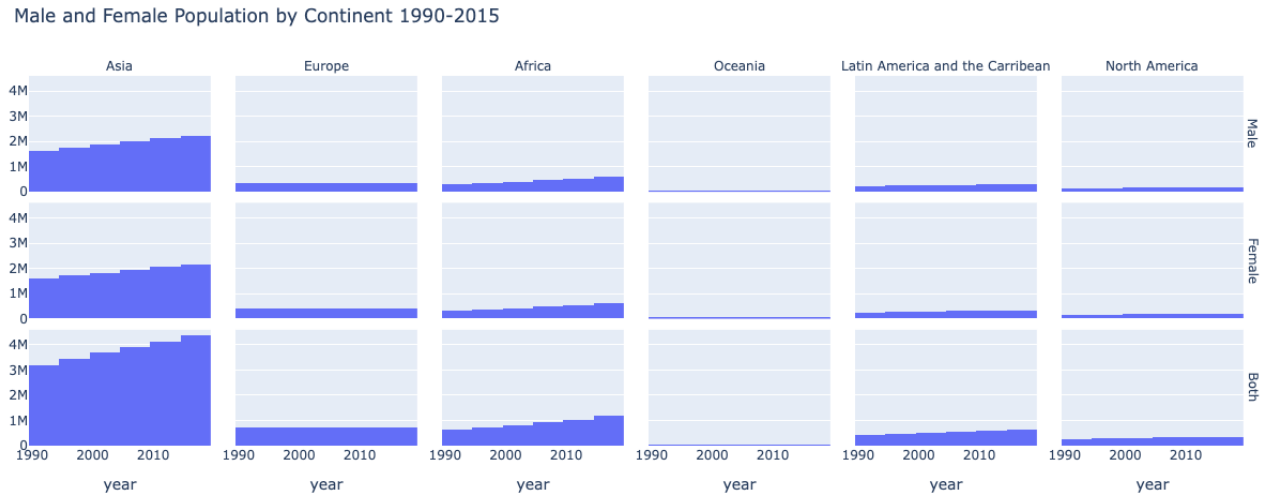


Fig 7. Male and Female Pop by Continent 1990-2015

By using multiples, we can see the overall distribution of the world's population across the continents. Not surprising, Asia is the top continent for population.

Table 3/DF3 – International Migrant Stock as a Percentage of the total population from 1990-2015

Table 3's observational units are quite rich compared to the last tables. The percentage of migrants in a country can display what countries are sought by migrants.

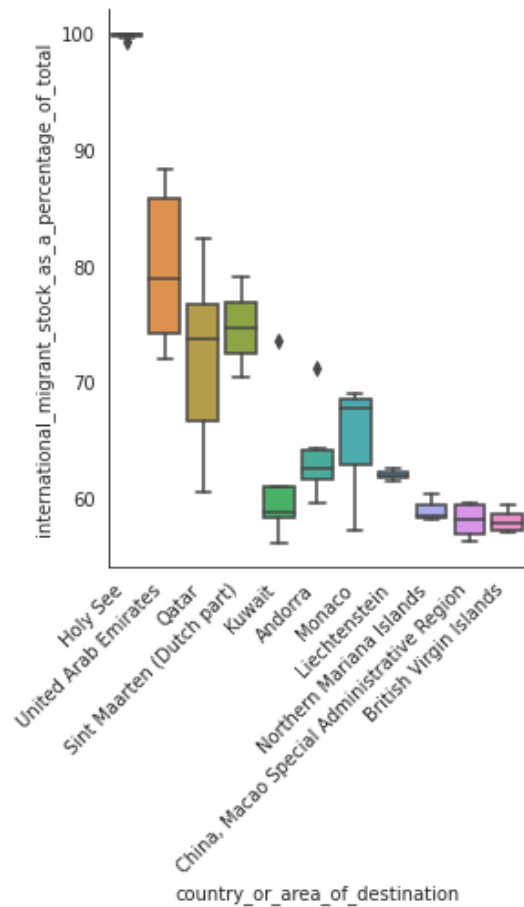


Fig 8 Distribution of Countries by Highest Percentage of Migrants in Population

The box chart shows an interesting result. The Holy See aka the Vatican does not per se have any citizens, rather it's a sovereign nation of the Catholic Church, therefore the Holy See would not have a citizenry. The UAE and Qatar are smaller countries that employ large amounts of migrant workers for construction. Monaco's migrant population has grown considerably as plotted by its tail. For this cursory look, the box chart is a good benchmark to work from.

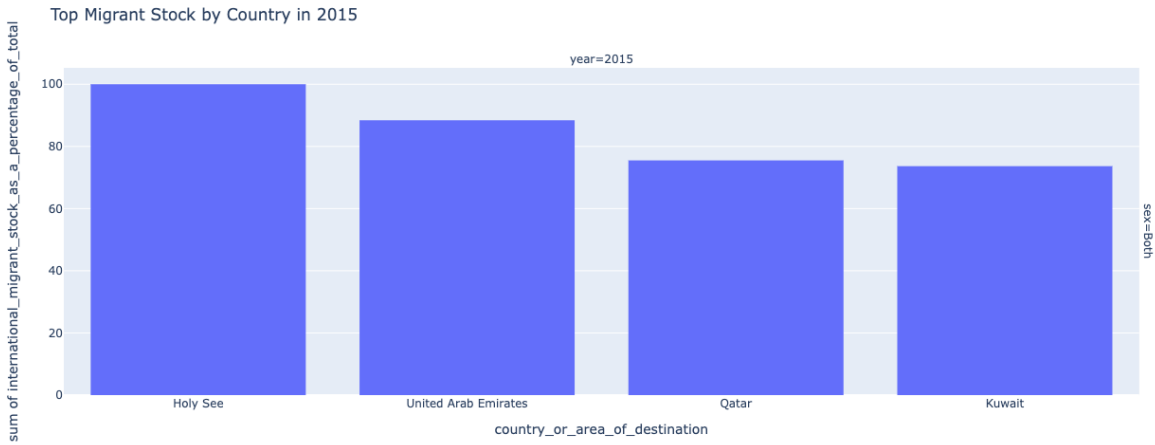


Fig 9 Top Migrant Stock by Country 2015

The figure above confirms our findings in the previous figure. We can see that the Holy See and UAE have the highest percentage of migrants to their populations, respectively. An interesting finding is the inclusion of Kuwait in 2015. This could be a result from migratory patterns as a result of the Iraq Wars. The box chart displays a more precise count that we can further investigate.

Male and Female Population by Percentage of Total 1990-2015

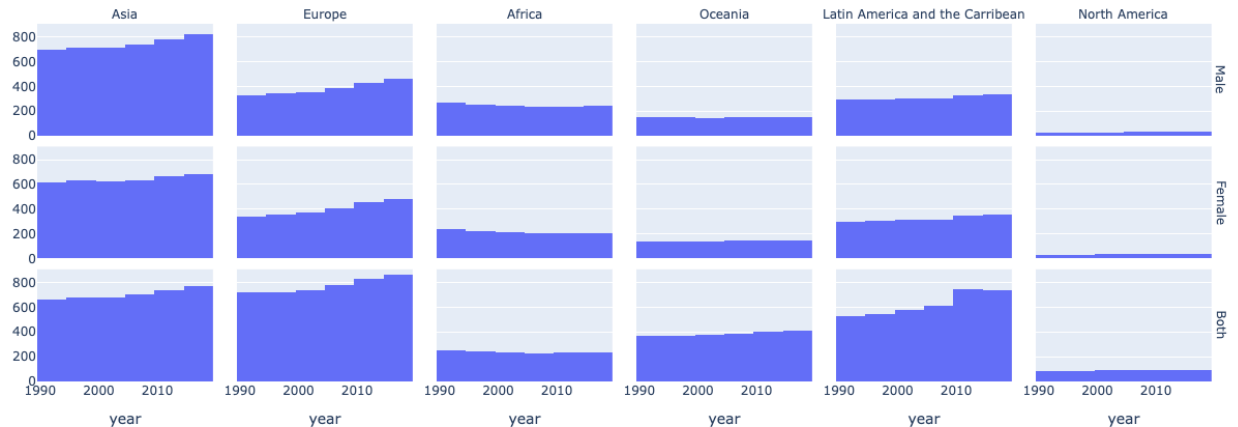


Fig 10 Male and Female Population by Percentage of Total 1990-2015

We can see from the above figure that Asia, Europe, and Latin America and the Caribbean, hold the largest migrant stock as a percentage to the total populations. There is a marked increase from 2000-2015. North America is relatively low migrants compared to the total populations.

Table 4/DF4 – Female Migrant Stock as a Percentage of the International Migrant Stock from 1990-2015

As noted from other tables, it will be interesting to see if Female migratory patterns differ from Males. There have been a couple of instances where female migrant stock eclipsed male.

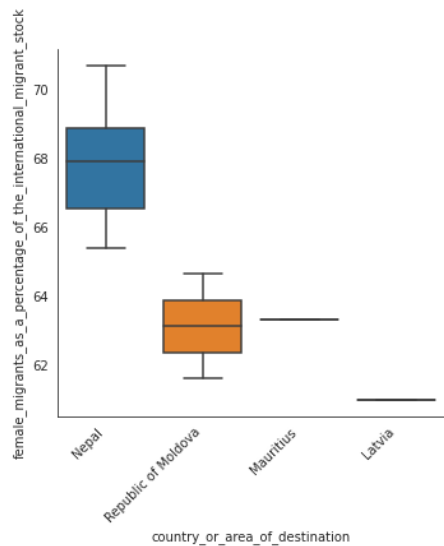


Fig 11. Top 4 Female Migrant Stock as a Percentage of the International Migrant Stock from 1990-2015

Nepal tops this list with considerable growth. This is due to Nepalese law stating that a female's citizenship is based on her father's citizenship (Pradhan 2020). This policy is distorting the percentage of female migrants for Nepal.

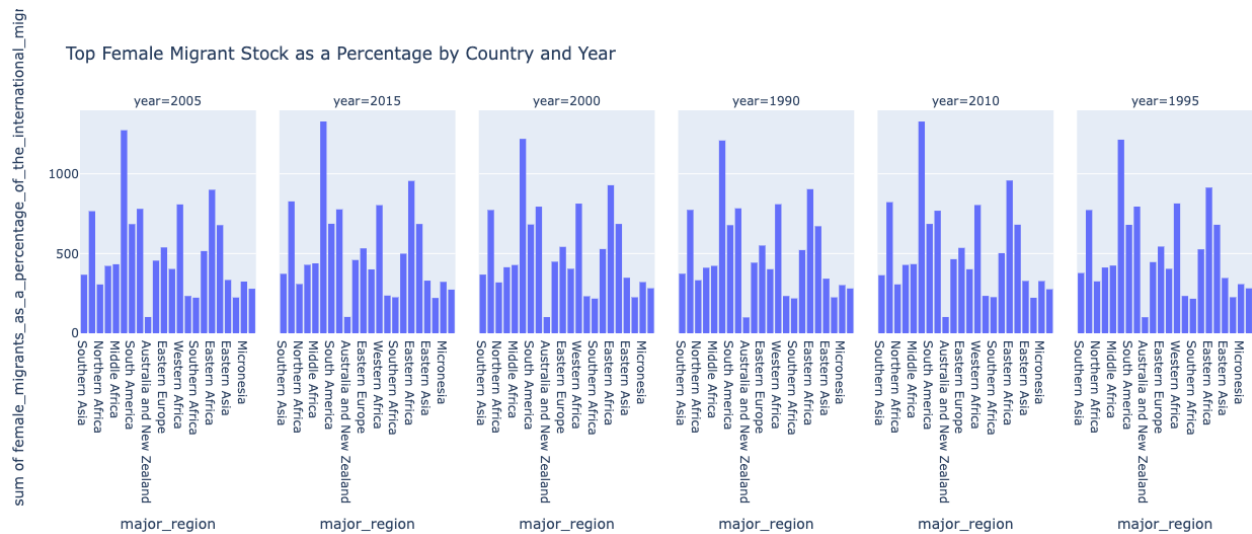


Fig 12 Top Female Migrant Stock as a Percentage by Country and year

From the figure above, the multiples showcase the surge in Female Migrant Stock by country. It's worth noting the prevalence of North America and the Caribbean in terms of Female Migrant Stock.

Table 5/DF5 – Female Migrant Stock as a Percentage of the International Migrant Stock from 1990-2015

Top Five Countries, annual rate of change by percentage of migrant stock, both sexes, from 2010-2015

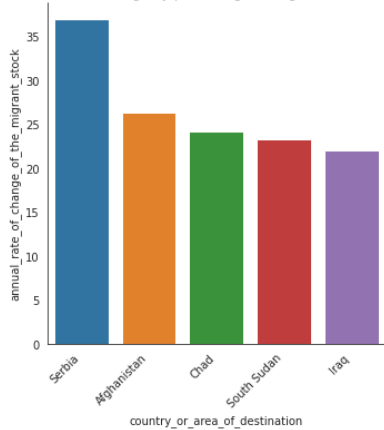


Fig 13. Top Five Countries, annual rate of change by percentage of migrant stock, both sexes 2010-2015

From the histogram above, we can see that Serbia, Afghanistan, Iraq, and Chad/South Sudan have the highest percentage of the migrant stock. This is sadly due to war. Chad accepted refugees from the civil conflict in Nigeria (“Chad appeals...” 2015).

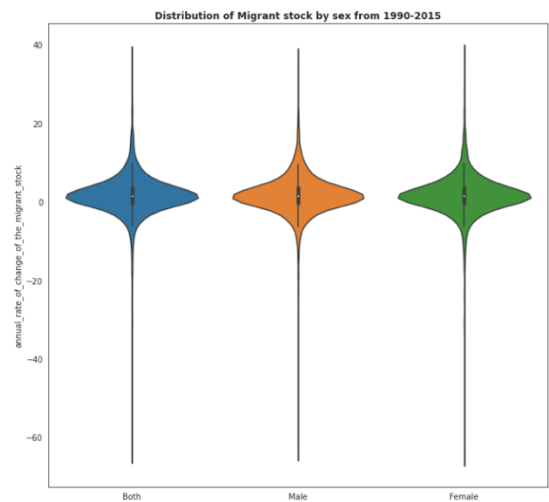


Fig. 14 Distribution of Migrant stock by sexes

The violin chart above shows the spread of the rate of change. We can see that the rate of change differs considerably with a -60 to 40 percentage. The median for both sexes show a minimal rate of change, hovering at about 5 percent. The violin chart can show these types of changes utilizing the whiskers, hinges, and median line.

Table 6A/DF6A – Estimated Refugee Stock at Midyear from 1990-2015

Top 20 Int. Migrant Stock Estimated Rate of Change from 1990-2015 by Country

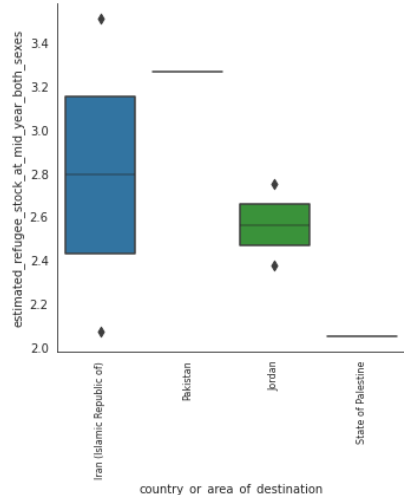


Fig. 15 Top 20 International Migrant Stock Estimated Rate of Change from 1990-2015 by Country

The box chart above shows countries with the highest estimated rate of change. Iran, Pakistan, Jordan, and Palestine share issues with conflict. Jordan and Pakistan are receiving countries for international migrants while Iran and Palestine contribute to the international migrant populations.

Table 6B/DF6B – Refugees as a percentage of the international migrant stock from 1990-2015

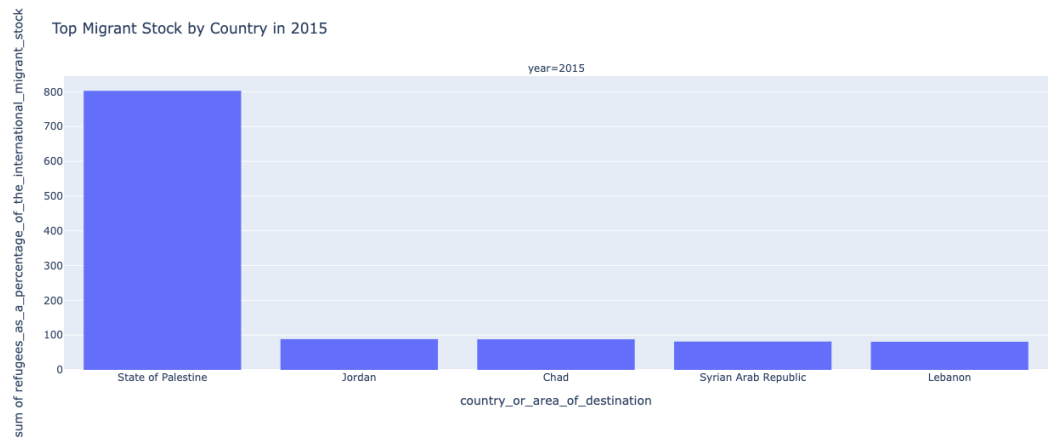


Fig 16 Top Migrant Stock by Country in 2015

By isolating the year, we can see that Palestine is an outlier by contributing to the international migrant stock. This could be due to political strife with Israel.

Table 6C/DF6C – Annual Rate of Change of Migrant Stock from 1990-2015

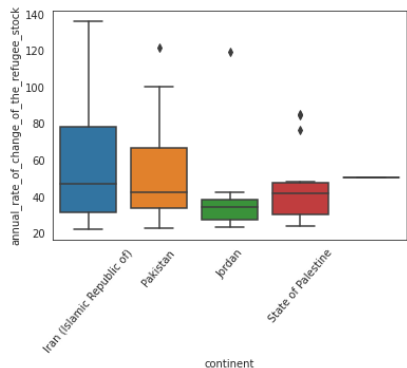


Fig. 17 Top 5 Countries with Annual Rate of Change of Migrant Stock from 1990-2015

This boxplot confirms our previous findings that Iran and Pakistan have the greatest annual rate of change of migrant stock.

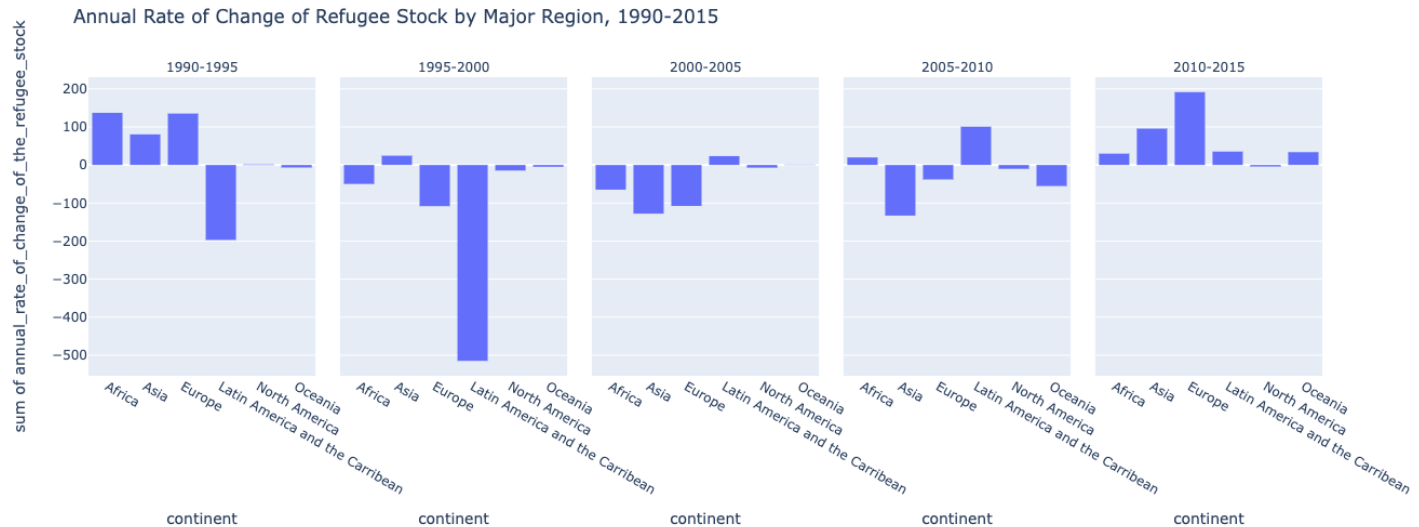


Fig. 18 Annual Rate of Change of Refugee Stock by Major Region, 1990-2015

The small multiple figures above display both the positive and negative rate of change by region. This gives an overall perspective of what areas contribute and accept the migrant stock.

Discussion

There are limitations to my coding ability that hindered the analysis. The use of overlays was a function that I couldn't figure out. I wanted to show the difference between tables like a country's total population versus migrant populations to show the gradual increase over time of migrants to a country. While I was able to calculate the range and mean of countries migrant stock, I was never able to plot these points.

My hope was to show the cumulative range and mean for all countries as a boxplot. I'd then create a side-by-side comparison of this boxplots to a single country. This would show the placement of the country against the overall trends for the year. I know Tukey stresses the importance of hinges and whiskers. I understand the value now in the EDA process. I was able to

see trends and patterns through whiskers, while hinges showed count tendencies, either positive or negative.

Conclusion

To answer my question above, the data set presented some unique answers for mass migration. There seems to be a tendency for migratory patterns to be triggered by human intervention other than wars. Nepal is a surprising case of gender inequality affecting citizenship. These distorted numbers show a higher influx of female migrants. The other was finding the correlation to world events like the fall of the USSR and rising migrant stock. These events trigger mass migratory movements that stretch over a decade.

References

- Chad Appeals for aid for refugees who fled Boko Haram attacks*. Reuters. (2015, January 8). Retrieved December 12, 2022, from <https://www.reuters.com/article/us-chad-nigeria-boko-haram-idUSKBN0KH1YE20150108>
- Pradhan, A. (2020, December 19). *Nepal's unequal Citizenship Law*. Nepali Times. Retrieved December 15, 2022, from <https://www.nepalitimes.com/banner/nepal-waits-for-an-unequal-citizenship-law/>